Distributed Source Coding

By

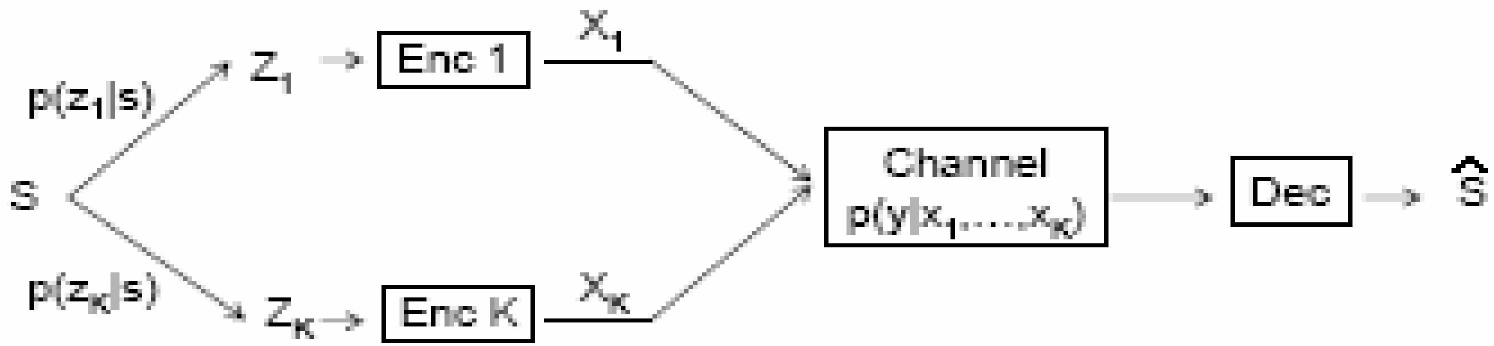
Vikrham Gowreesunker

Overview

- Distributed Source Coding
  - 'Compression of multiple correlated sensor outputs that do not communicate with each other. The sensors send their compressed outputs for a centralized joint decoding'
- Lossless source coding of discrete source w/ side information at decoder as a case Slepian-Wolf Coding
- The Wyner-Ziv Coding Scheme

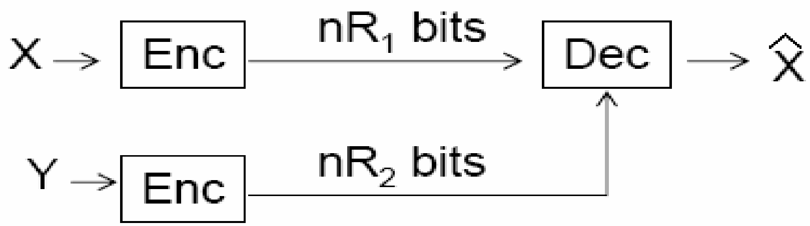
Graphics were borrowed from Dr.Jindal's EE8510 class notes and "Distributed Source Coding for sensor networks" by Xiong, Liveris and Cheng

Sensor Network Model



- S: source (random variable)
- $Z_1,\ldots,Z_K$: sensor observations
- $X_1,\ldots,X_k$: sensor transmissions
- Y: channel output
- $\hat{S}$: estimate of source

Source Coding with Side Info



- Want to decode X losslessly using Y only as side information
- (R_1,R_2) achievable iff (Wyner) [5]:

$$R_1 \geq H(X \mid U)$$

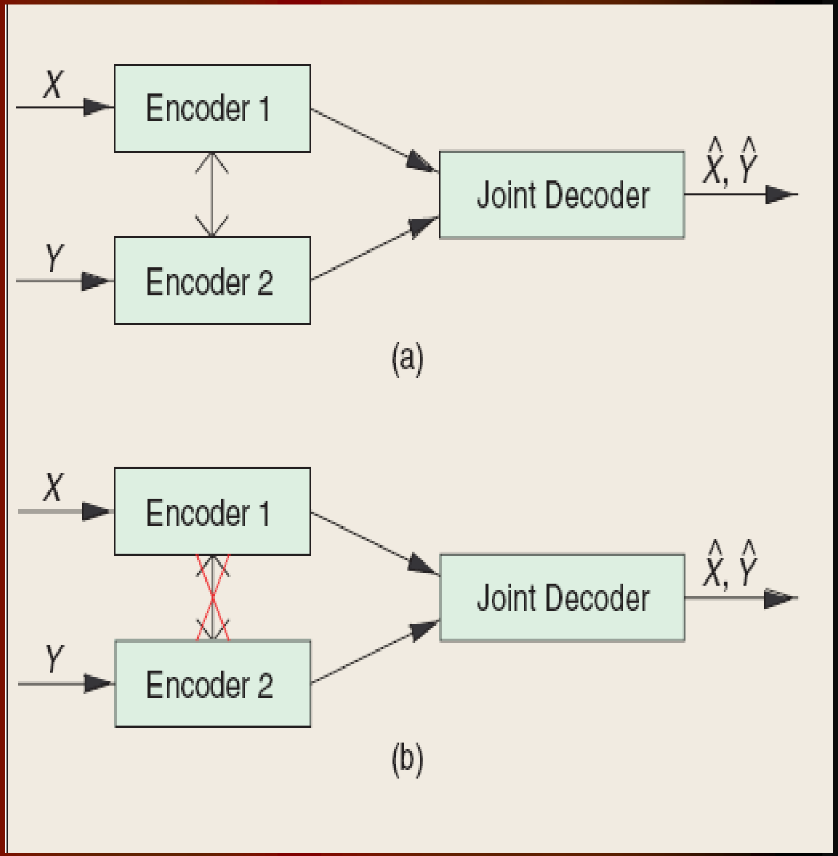
$$R_2 \geq I(Y;U)$$

for some p(u|y)

- Achievability: Cover Y with reconstruction r.v. U, use random binning on X and Markov lemma to find typical (X^n,U^n) pair in bin
- Markov Lemma: If X->Y->Z, (x^n,y^n) typical, and z^n chosen iid $p(z^n|y^n)$, then (x^n,y^n,Z^n) typical with nearly probability one
  - By definition X->Y->U. Since (x^n, y^n) typcial with high probability, also have (x^n,y^n,u^n) typical, which implies (x^n,u^n) typical

Slepian-Wolf Coding (SWC)

- (Xi,Yi), i.i.d from a pair of correlated discrete r.v X and Y.
- Part (a)
  - For Joint Encoding, a rate of H(X,Y) is sufficient
- Part (b)
  - For distributed (non cooperative) encoding, slepian-wolf showed that a rate of H(X,Y) is still sufficient



Slepian-Wolf Coding (SWC) cont..

- Proof: Uses the concept of random binning. This involves partitioning the space of all possible outcomes of a random source into disjoint sets or *bins*. [slepian-wolf]

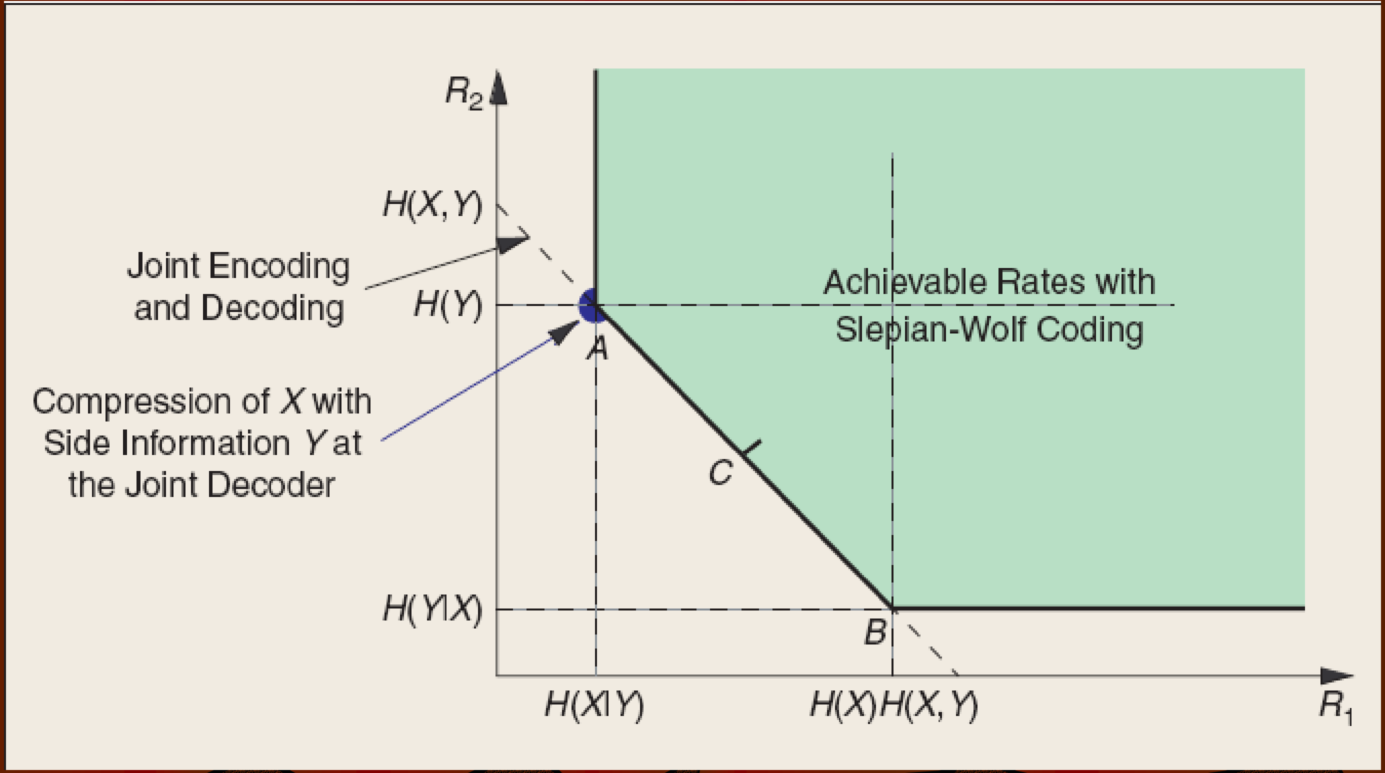
- Generalization of achievability [Cover] to include:

  Arbitrary ergodic processes, countably infinite alphabets and arbitrary number of sources.

- In practice, random binning is non constructive and methods such as pseudo-random binning and algebraic binning could be used. However, a more appealing strategy was brought forward by Wyner.

Slepian-Wolf Rate Regions

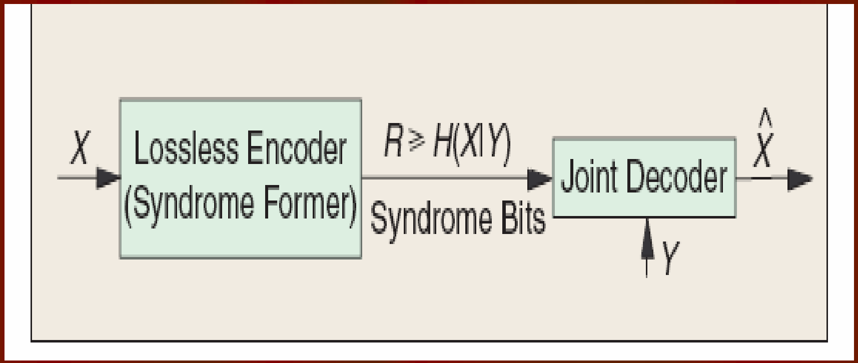
- For 2 sources, we have the following rate region.



SWC as Source Coding with SI at Decoder

- The point A requires a rate of R1+R2 = H(X,Y) = H(X/Y)+H(Y). This is the same as source coding with side info at decoder.

- If A is achievable, B can be achieved by inverting X and Y. All rates in between can be achieved by time sharing.



Asymmetric Coding

- In 1974, Wyner drew the parallel between asymmetric coding and the well studied channel coding problem. He suggested the use of linear block codes.

- Wyner's Scheme
  - Binary Symmetric Sources & Hamming distance measure.
  - (n,k) binary block code => 2^(n-k) syndromes, each corresponding to a set of 2^k binary words of length n.
  - Each bin is a *coset code.*
  - *n-bits are mapped into n-k bits index, achieving a compression*
  - *ratio of n : (n-k).*

- However, this approach has began to appear in practical codes only recently. [Pradhan and Ramchandran 2003].

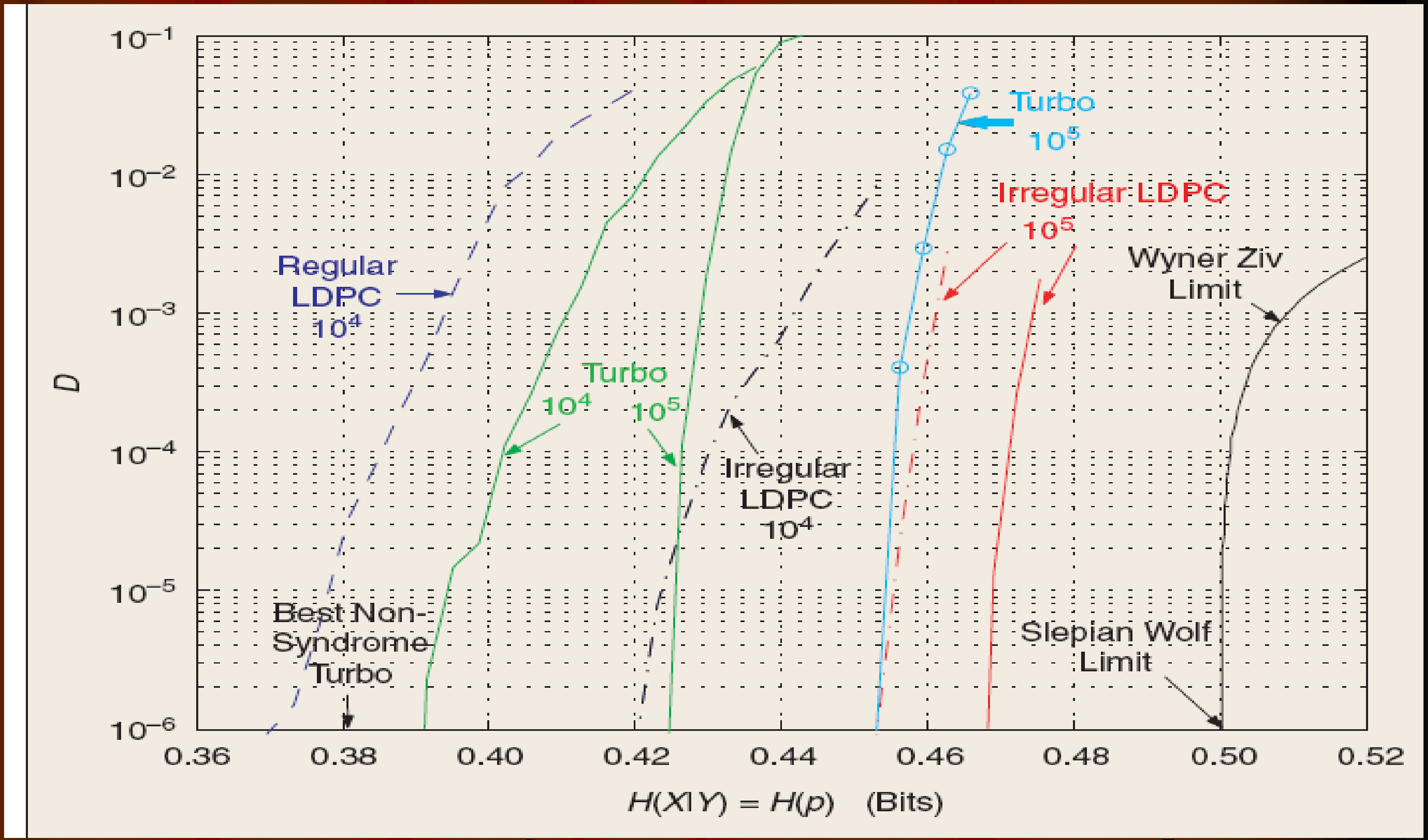
Asymmetric Coding cont..

- Wyner’s Syndrome concept can be extended to all binary linear codes and near-capacity channel codes such as Turbo and Low Density Parity Check (LCPD) codes.
- Turbo and LCPD codes have been shown to approach the slepian-wolf limit.
- In practice, linear channel code rates depends of the correlation model between X and Y.
- Next we look at the binary symmetric correlation model

SWC for Binary Symmetric Correlation

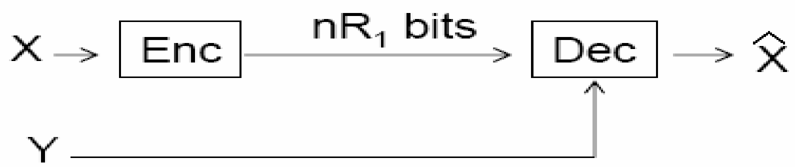
- The first practical designs borrowed ideas from Turbo codes and Parity Bits for this model.
- Liveris et al. followed the Wyner’s Scheme with Turbo/LDPC codes. The perfomance was very close to SW limits.
- Simulations results from Liveris et al are shown in the following figure

Liveris et al Results



Wyner-Ziv Coding (WZC)

Rate-Distortion with Side Info



- Want to lossy decoding of X using Y as side information (Wyner-Ziv) [6]

$$R_Y(D) = \min_{\substack{p(u \mid x),\, f(y,u) \\ E[d(x, f(y,u))] \le D}} I(X;U) - I(Y;U)$$

- Generally is <u>strict</u> rate loss relative to cooperative encoder (i.e. encoder who knows X and Y)
- Achievability: Generate $2^{nI(X;U)}$ codewords iid p(u) and randomly place in 2^{nR} bins. Choose codeword typical with x^n and send bin index. Decoder finds (u^n, y^n) typical (Markov lemma) in correct bin. Given correct u^n, computes f(u,y) to estimate X.

Rate Loss and Wyner-Ziv

- For zero mean and jointly gaussian X and Y and MSE, Wyner-Ziv Coding has no rate loss when compared to joint encoding.

- Pradhan et al, shows that this is also valid for the more general case of $X = Y + Z$ with Z gaussian and X and Y arbitrarily distributed.

WZC as source-channel coding problem

- Source coding is employed to quantized X.
- Furthermore, the quantized version is still correlated to the side information. Hence rate can be reduced using Slepian-Wolf coding.
- SWC can essentially be viewed as channel coding. Hence the joint source-channel coding.
- In practice, we use both
  1. A Source Code [e.g. Trellis Coded Quantization] that can achieve granular gain
  2. A Channel Code [e.g. turbo and LDPC] that approach the slepian-wolf
- DISCUS [1999] is the first major work on Wyner-Ziv code design. Several other designs have since been proposed. Among the most notable one is the Slepian-Wolf coded quantization(SWCQ). This combines source coding and channel coding for algebraic binning.

Further Work

- Theoretical limits for *lossy* Distributed Source Coding for multiple sources.
- Practical Codes for CEO problem [currently limited to special cases]
- Cross-Layer design aspect of DSC.